

Textures Dynamiques: état de l’art, modélisation, applications

Sloven DUBOIS^{1,2}, Renaud PÉTERI¹, Michel MÉNARD²

¹Laboratoire MIA, Av. Michel Crépeau, 17000 La Rochelle, France

²Laboratoire L3i, Av. Michel Crépeau, 17000 La Rochelle, France

sloven.dubois01@univ-lr.fr, renaud.peteri@univ-lr.fr, michel.menard@univ-lr.fr

Résumé – Le contexte de nos travaux de recherche porte sur l’étude de mouvements répétitifs en temps et en espace, appelés Textures Dynamiques (TD), et sur la manière d’en extraire des caractéristiques spatiales et temporelles par des décompositions multi-échelles adaptées. Cet article est composé de deux parties. La première aborde rapidement un état de l’art de la TD qui permet ensuite d’établir une taxonomie ainsi qu’une définition de ce phénomène complexe. Nous proposons également un modèle formel. Ensuite dans une deuxième partie, nous abordons une application de recherche de TD par le contenu.

Abstract – The context of our research involves the study of repetitive movements both time and space, known as dynamic textures (DT), and the extraction of spatial and temporal characteristics by multi-scale decompositions. This article is composed of two parts. The first quickly exposed a state of the art of DT, a taxonomy and a definition of this complex phenomenon are established. After a formal model is proposed. In second part, applications of this formal model is illustrated.

1 Introduction

Un drapeau dans le vent, un champ d’herbe ondulant, les vagues de la mer, la surface d’un lac, le mouvement de la forêt, la fumée, le feu, *etc.*, sont autant d’exemples de Textures Dynamiques (TD) présentés dans la littérature. Elles représentent un thème de recherche très actuel, le nombre de publications portant sur la caractérisation de TD est en forte augmentation ces dernières années. Les premières publications sur ce domaine apparaissent en 1992 avec l’article de Nelson et Polana [9], dans lequel les auteurs parlent de textures temporelles, et définissent un premier cadre formel d’étude. Par la suite, à partir de l’année 2002, le nombre de publications proposant des méthodes pour les caractériser augmente considérablement.

L’émergence de ce sujet de recherche peut s’expliquer, d’une part, par la démocratisation du support vidéo et des performances des ordinateurs pouvant les traiter, et d’autre part, par un vaste champ applicatif : indexation vidéo, vidéo surveillance, segmentation spatio-temporelle, soustraction de fonds dynamiques, suivi de TD, synthèse de vidéos, *etc.*

Le parcours de la littérature permet de constater que les phénomènes complexes que nous nommons TD, sont identifiés sous des termes différents selon les auteurs : *temporal textures* [9], *spatio-temporal textures* [10], *dynamic textures* [12], *time-varying textures* [1]. Cette variabilité est bien sûr due à la diversité des définitions sous-jacentes proposées par les auteurs. Un recensement des principales publications sur le thème des TD permet d’établir que 75% d’entre elles ne précisent pas de définition de ce phénomène. Le domaine de l’image se heurte à des difficultés importantes pour définir de manière formelle une texture spatiale, l’ajout de la dimension temporelle com-

plique encore la tâche.

Dans cet article, nous proposons donc une taxonomie des TD et un modèle pour certaines d’entre elles, puis nous présentons une application de recherche de TD par le contenu.

2 Taxonomie des textures dynamiques

La figure 1 présente une taxonomie des TD. Elle se caractérise de la manière suivante :

- une séquence d’images peut contenir des composantes textures statique et/ou dynamique. Dans la vidéo présentée dans la figure 1, le pont, les cailloux et le lierre sont des éléments texturés statiques alors que l’eau et l’arbre sont en mouvement ce qui induit une TD.

- la composante TD d’une séquence d’images peut contenir plusieurs TD se différenciant par leur nature. Ici, la composante texture vidéo est constituée de deux TD distinctes : l’arbre et le ruisseau. Ces différentes TD peuvent bien sûr interférer, ce qui n’est pas le cas ici. Dans le cadre d’une séquence avec du feu, les oscillations des flammes, les volutes de fumée en avant plan de l’oscillation d’un champ d’herbe interfèrent pour former une composante TD complexe.

- quand on observe de nombreuses scènes, nous constatons que les TD ne sont pas toutes de même nature : naturelles (issues de la nature), artificielles (créées par l’être humain) et synthétiques (produites par un ordinateur). Une TD se caractérise par trois facteurs :

- une structure texturée (rigide ou déformable).
- un mouvement induit par une force que subit la structure texturée ou par le déplacement de la caméra : ce mouve-

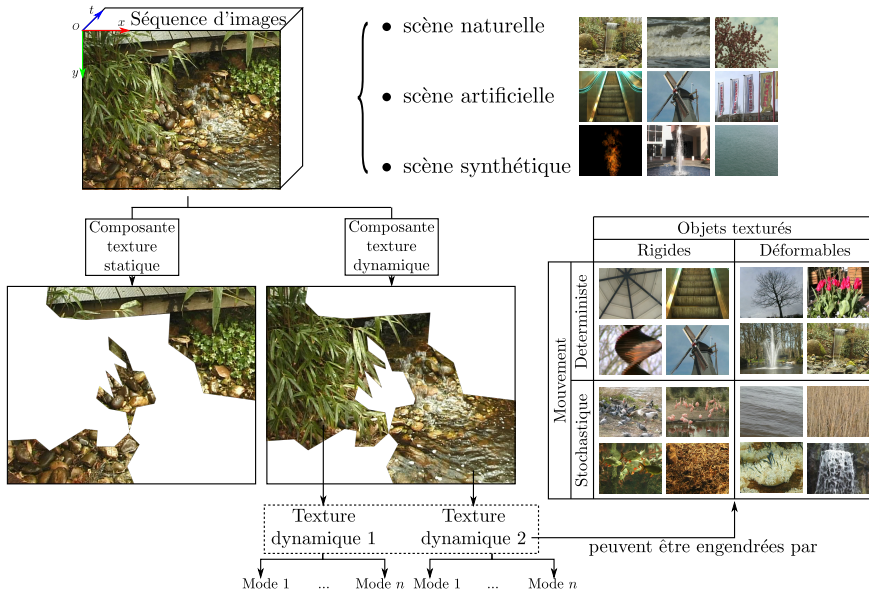


FIG. 1 – Taxonomie des textures dynamiques.

ment peut être d'apparence déterministe ou stochastique. Celui-ci peut être de plusieurs natures : force interne (moteur d'un escalator ou d'une voiture), force externe (le vent dans les ailes d'un moulin) ou induite par la caméra (translation de la caméra).

- des changements de condition d'acquisition (éclairage, illuminant, etc). Ces modifications induisent une modification apparente de la texture et donc l'apparition d'une TD.

Les ailes d'un moulin présentant un mouvement de rotation, les marches d'un escalator défilant, le trafic routier, sont des exemples de TD issues d'une structure texturée rigide possédant un mouvement déterministe. Par contre, un banc de poisson, une colonie de fourmis, sont également des structures texturées rigides mais qui ont un mouvement stochastique.

Les TD constituées de structures texturées déformables et présentant des mouvements stochastiques sont par exemple une cascade d'eau avec de multiples remous, une anémone ballottée par le courant, etc. Par contre, un arbre ou des fleurs oscillants sous un vent constant, une cascade sans remous, sont des exemples de TD issues de structures déformables et présentant un mouvement déterministe. Toutefois, selon l'échelle d'observation, certains de ces phénomènes peuvent être vus comme des structures rigides animées par un mouvement déterministe.

- l'observation de certaines TD permet de constater qu'elles peuvent être la composition de plusieurs phénomènes. Prenons l'exemple de la vidéo de la surface de la mer (cf. figure 2), on observe des risées (mouvement haute fréquence) portées par les vagues (mouvement basse fréquence). Ces composantes appartiennent à la même TD et se définissent comme des modes [11] pouvant se superposer.

L'ensemble de ces réflexions nous conduit à la définition :

Une séquence d'images naturelle, artificielle ou synthétique peut contenir une composante texture statique et/ou une com-

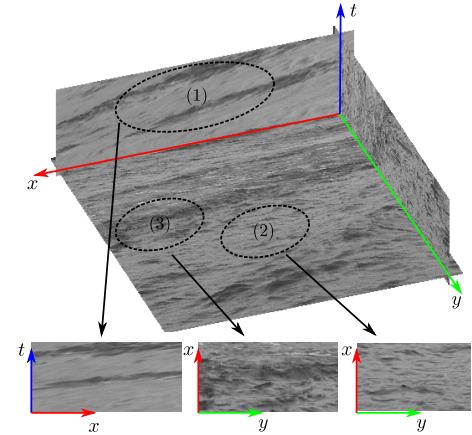


FIG. 2 – Coupes spatio-temporelles d'une texture dynamique. On peut observer l'onde porteuse (1), les phénomènes locaux (2) et l'interférence entre les deux (3).

posante TD. Cette dernière est composée d'au moins une TD. Une TD est une structure texturée qui peut être rigide ou déformable. Cette structure possède un mouvement induit par une force qu'elle subit ou par le déplacement de la caméra. Ce mouvement, qu'il soit issu d'une force interne, d'une force externe ou induit par la caméra, peut être déterministe ou stochastique. Les TD sont constituées de modes, pouvant se superposer, caractérisés par des phénomènes spatiaux et temporels répétitifs.

Afin de parfaire cette taxonomie, un modèle formel, inspiré de la synthèse d'images [6] et de l'observation de la base DynTex [11], est maintenant présenté.

3 Modélisation des Textures Dynamiques

L'observation des TD de DynTex permet d'établir que :

- une TD peut posséder un support spatio-temporel différent de celui de la vidéo.
- les supports des différentes TD peuvent être disjoints. Certaines TD sont en effet partiellement transparentes (l'eau, la fumée). D'autres sont composées d'éléments discrets et peu denses laissant apparaître l'arrière-plan (les herbes, les arbres).
- différents phénomènes (cf. figure 2) cohabitent dans une TD :
 - une onde porteuse (1) : onde à grande échelle se propageant au cours du temps. Ce phénomène, comme son nom l'indique, "porte" l'information.
 - des phénomènes localement oscillants (2) : de petites oscillations bien localisées dans la séquence d'images, de support réduit, et de fréquences élevées, apparaissant de manière stochastique.
 - la superposition des deux phénomènes (3) : à certains moments des séquences, les deux phénomènes interfèrent.

Les TD sont ainsi perçues comme la superposition de plusieurs phénomènes physiques. Ainsi la composante TD d'une vidéo T_V se définit comme la somme de $N \in \mathbb{N}^*$ Textures Dynamiques $\Upsilon_i \in \mathbb{R}$, chacune de support spatio-temporel Ω_i pouvant être disjoint : $T_V(\mathbf{x}) = \sum_{i=1}^N \Upsilon_i^{\Omega_i}(\mathbf{x})$ où $\mathbf{x} = (x, y, t)^T$ représente les coordonnées d'un voxel dans le cube vidéo.

Texture Dynamique Υ_i : D'après l'observation de la base DynTex et de résultats obtenus en synthèse d'images [6], Υ_i peut s'écrire comme la composition d'une onde se propageant à grande échelle et de phénomènes localement oscillants ; elle suit alors le modèle suivant :

$$\forall i, \quad \Upsilon_i^{\Omega_i}(\mathbf{x}) = \mathcal{P}_i(\mathbf{x}) + \mathcal{L}_i(\mathbf{x}) \quad (1)$$

où \mathcal{P}_i et \mathcal{L}_i sont deux fonctions décrivant respectivement l'onde porteuse et les phénomènes locaux d'une TD Υ_i .

Le modèle ci-dessus formalise les classes de TD suivantes :

- les structures déformables possédant un mouvement stochastique ou déterministe comme les écoulements de fluides (lac, mer, ruisseau), les oscillations engendrées par le vent (l'herbe, les arbres, les drapeaux), la propagation de fumée, *etc.*
- les structures rigides possédant un mouvement déterministe comme les escalators, les ailes d'un moulin, *etc.*
- une partie des structures rigides possédant un mouvement stochastique comme un banc de poisson, un essaim d'abeilles, *etc.*

L'onde porteuse \mathcal{P}_i : Il s'agit d'un phénomène complexe. Son support spatio-temporel dépend bien sûr de la séquence. L'onde porteuse se caractérise par sa vitesse de propagation, sa direction et son degré de stationnarité. De plus, il s'agit d'un phénomène périodique en temps et en espace. L'onde porteuse \mathcal{P}_i d'une TD Υ_i s'écrit comme une somme d'exponentielles complexes spatio-temporelles :

$$\mathcal{P}_i(\mathbf{x}) = \sum_{p_n \in \mathcal{P}_i} A_{p_n}(\mathbf{x}) \operatorname{Re} \left(e^{j(\omega_{p_n} \cdot \mathbf{x} + \psi_{p_n})} \right) \quad (2)$$

d'amplitude $A_{p_n} \in \mathbb{R}^{+*}$, de pulsation $\omega_{p_n} \in \mathbb{R}^3$ et de déphasage $\psi_{p_n} \in \mathbb{R}$ propres à chacune d'elles. Ce phénomène, comme son nom l'indique, propage l'information de la texture et plus précisément les phénomènes locaux.

Les phénomènes locaux \mathcal{L}_i : Il s'agit de phénomènes localement oscillants à la fois en espace et en temps. Le modèle prend en compte cette localisation spatio-temporelle grâce à un noyau gaussien spatio-temporel dont les paramètres prennent en compte la rapidité d'apparition/disparition des phénomènes ainsi que leur support spatial. Les phénomènes locaux \mathcal{L}_i d'une TD Υ_i se définissent par :

$$\mathcal{L}_i(\mathbf{x}) = \sum_{\ell \in \mathcal{L}_i} \mathcal{N}_G(\mu_\ell, \Sigma_\ell)(\mathbf{x}) \sum_{\ell_k \in \mathcal{L}_i} A_{\ell_k}(\mathbf{x}) \operatorname{Re} \left(e^{j(\omega_{\ell_k} \cdot \mathbf{x} + \psi_{\ell_k})} \right) \quad (3)$$

avec pour un phénomène local ℓ particulier, $\mathcal{N}_G(\mu_\ell, \Sigma_\ell)$ un noyau gaussien associé qui décrit sa localisation spatio-temporelle et $A_{\ell_k} \in \mathbb{R}^{+*}$, $\omega_{\ell_k} \in \mathbb{R}^3$ et $\psi_{\ell_k} \in \mathbb{R}$ représentent respectivement l'amplitude, la pulsation et le déphasage.

Ce modèle formel ramène l'analyse des TD sous la forme

d'une décomposition de celles-ci en phénomènes locaux oscillants et en ondes porteuses. Dans la littérature, de nombreuses approches permettent la décomposition d'un signal en différentes composantes : décomposition de Meyer [8], séparation aveugle de source [2], décomposition de Wold [7], *etc.*

En raison des nombreuses bases utilisables dans le dictionnaire d'analyse, notre choix s'est porté sur l'analyse en composantes morphologiques que nous avons étendue à la dimension temporelle en utilisant une nouvelle stratégie de seuillage afin de diminuer les temps de calcul [4, 5]. Ainsi, la transformée en curvelets 2D+T est utilisée pour caractériser les ondes porteuses et la transformée en cosinus locale 2D+T pour les phénomènes locaux.

La figure 3 présente une séquence d'images correspondant à la surface de la mer au bord d'une plage. Il s'agit d'une TD dont les propriétés sont proches de celles de notre modèle : elle est constituée d'une onde porteuse (les rouleaux des vagues) et de phénomènes locaux (l'écume). Dans les résultats de la décomposition, illustrés sur la figure 3, on retrouve les caractéristiques de notre modèle dans chacune des composantes. On observe, dans les représentations surfaciques, que la partie géométrique met en évidence le rouleau des vagues qui n'est plus perturbé par les phénomènes locaux oscillants. Cette simplification permet une meilleure compréhension du phénomène complexe de la TD.

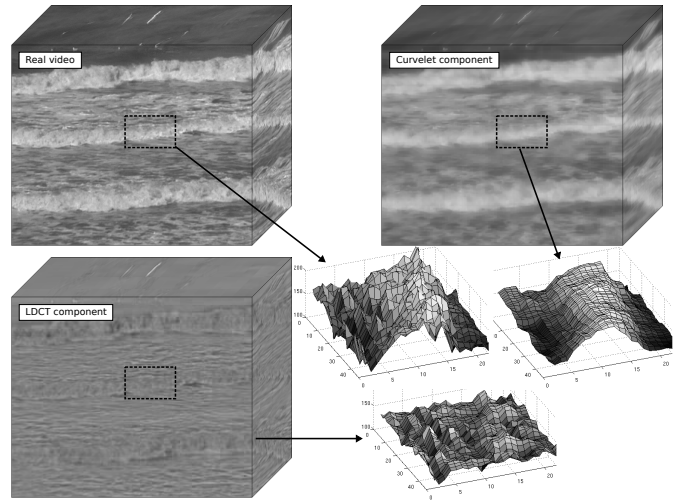


FIG. 3 – Décomposition de textures dynamiques

4 Recherche de TD par le contenu

Ce modèle a été utilisé dans diverses applications [3] : simplification de TD, segmentation, indexation, estimation du mouvement globale, recherche par le contenu. Nous présentons ici la recherche de TD par le contenu.

L'objectif est de présenter une requête sur la base DynTex sous la forme d'une séquence d'images et d'observer celles retenues par le moteur de recherche.

La figure 4 présente les résultats de quatre requêtes. La première colonne présente la séquence d'images requête, suivie

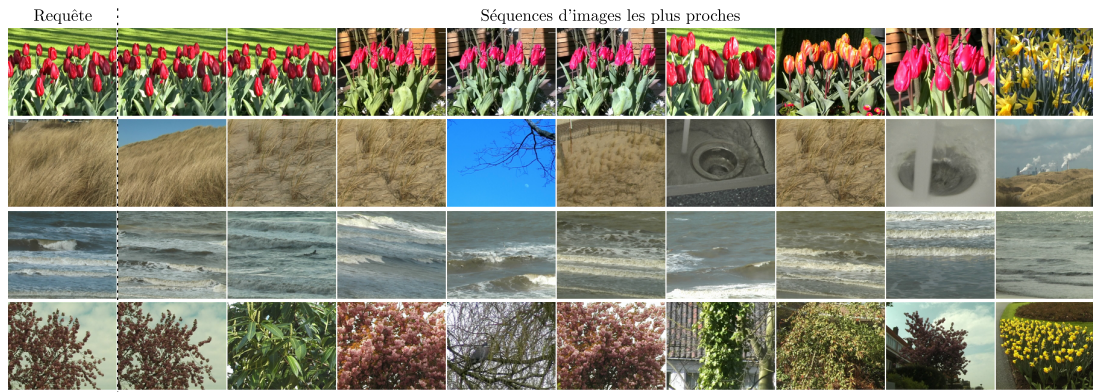


FIG. 4 – Recherche par le contenu de textures dynamiques

pour chacune d'elle, des 9 séquences les plus proches. Nous nous appuyons ici sur une mesure de distance entre la signature (calculée sur les composantes issues de l'analyse en composantes morphologiques [3]) de la séquence requête et celles des séquences de la base sans aucune autre information ajoutée (couleur, structure, etc).

Différentes observations peuvent être effectuées :

- Fleur : les 9 séquences d'images les plus proches sont pertinentes et représentatives de la requête. Malgré un contenu identique, la sixième séquence n'apparaît pas dans les premiers résultats à cause du changement d'échelle, pris en compte dans notre décomposition. Le changement de couleur ou d'arrière plan n'interfère pas dans le calcul du vecteur de caractéristiques.
- Herbe : pour cette requête, les premières séquences trouvées sont proches de la vidéo requête. Par la suite nous notons plusieurs aberrations. La présence des deux séquences d'images "Évier" traduit le fait que les descripteurs considérés sont relativement proches dans l'espace des caractéristiques.
- Mer : comme dans le cadre des fleurs, les 9 séquences résultats sont proches de la séquence requête. On constate en effet que l'ensemble des TD sélectionnées se ressemble du point de vue des caractéristiques spatio-temporelles.
- Arbre : les résultats obtenus par la recherche correspondent à des séquences d'images de feuillage exceptées pour deux d'entre elles : la quatrième et la neuvième séquence. Cependant, la structure de branches sans feuillage, et le massif de fleurs possèdent des caractéristiques spatio-temporelles proches de la séquence requête. Nous observons également que l'arbre de la séquence d'images requête se retrouve présent à différentes échelles dans les séquences résultats.

5 Conclusions et perspectives

Dans la première partie de cet article, nous avons établi une taxonomie des TD permettant ainsi de décrire ce phénomène complexe. Ensuite, un modèle formel de TD est présenté. Celui-ci définit une classe importante de TD comme étant la somme d'une ou plusieurs ondes porteuses et de phénomènes locaux. En s'appuyant sur ce modèle, une application de recherche de TD par le contenu est ensuite proposée et discutée.

L'étude des TD est encore un domaine de recherche jeune, et nous pensons qu'il continuera à croître fortement dans les prochaines années. Prendre en compte à la fois la complexité et la grande variabilité de ces TD est en effet un défi majeur. Le développement d'outils avancés permettant de progresser dans la compréhension et l'analyse de ces TD est un enjeu important : les domaines applicatifs où interviennent les TD étant en effet abondants et sont amenés à croître avec le futur.

Références

- [1] Z. Bar-Joseph, R. El-Yaniv, D. Lischinski, and M. Werman. Texture mixing and texture movie synthesis using statistical learning. *IEEE Trans. on VCG*, 7 :120–135, 2001.
- [2] J. Bobin, J.L. Starck, Y. Moudden, and J.M. Fadili. Blind source separation : the sparsity revolution. *AIEP*, pages 221–298, 2008.
- [3] S. Dubois. Décomposition spatio-temporelles pour l'étude des textures dynamiques : Contribution à l'indexation vidéo, Thèse de l'Université de La Rochelle, 200 pages, 2010.
- [4] S. Dubois, R. Péteri, and M. Ménard. Decomposition of dynamic textures using morphological component analysis : A new adaptative strategy. In *ICPR*, 2010.
- [5] S. Dubois, R. Péteri, and M. Ménard. Decomposition of dynamic textures using morphological component analysis. *IEEE Trans. on CSVT*, DOI :10.1109/TCSVT.2011.2159430, 2011.
- [6] M. Finch. *GPU Gems : Programming Techniques, Tips, and Tricks for Real-Time Graphics, Chap.1*. 2004.
- [7] J.M. Francos, A.Z. Meiri, and B. Porat. A unified texture model based on a 2-d wold-like decomposition. *IEEE Trans. on SP*, 41 :2665–2678, 1993.
- [8] Y. Meyer. *Oscillating Patterns in Image Processing and Non-linear Evolution Equations*. American Mathematical Society, Boston, MA, USA, 2001.
- [9] R.C. Nelson and R. Polana. Qualitative recognition of motion using temporal texture. *CVIU*, 56 :78–89, 1992.
- [10] C.H Peh and L.F. Cheong. Exploring video content in extended spatio-temporal textures. In *Eu. Wksp. on CBMI*, 1999.
- [11] R. Péteri, S. Fazekas, and M.J. Huiskes. Dyntex : A comprehensive database of dynamic textures. *PRL*, 31 :1627–1632, 2010.
- [12] P. Saisan, G. Doretto, Y.N. Wu, and S. Soatto. Dynamic texture recognition. In *CVPR*, pages 58–63, 2001.